Iranian Journal of Animal Science Research
Homepage: http://ijasr.um.ac.ir

Research Article
Vol. 15, No.4, 2023, p. 585-597

Comparison of Different Genomic Relationship Matrices for Multibreed Weighted Single Step GWAS

Seyed Mostafa Mazloom[1], Mohammad Mahdi Shariati[2]*

1- Ph.D. Candidate, Department of Animal Science, Ferdowsi University of Mashhad, Iran.
2- Assistant Professor, Department of Animal Science, Ferdowsi University of Mashhad, Iran.
*Corresponding Author's Email: mm.shariati@um.ac.ir

Received: 16-10-2022 Revised: 17-04-2023 Accepted: 21-05-2023 Available Online: 07-06-2022	**How to cite this article:** Mazloom, S. M., & Shariati, M. M. (2023). Comparison of different genomic relationship matrices for multibreed weighted single step GWAS. *Iranian Journal of Animal Science Research*, *15*(4), 585-597. (in Persian with English abstract). https://doi.org/10.22067/ijasr.2023.79619.1108

Introduction: Genomic best linear unbiased prediction (GBLUP) and Bayesian methods are used for genomic selection (GS) and genome wide association study (GWAS) in animal and plant breeding. The main objective of GWAS is detection of quantitative trait loci (QTL) that is affecting trait. The GBLUP assumes equal variance for all markers but the Bayesian methods assumes specific variance for each marker. When trait are affecting by major QTLs, Bayesian methods have the benefit of marker selection. In weighted GBLUP (WGBLUP), disparate variance of marker specific weights for weighting of all markers are used. If only a deduction of animals is genotyped, single-step WGBLUP (WssGBLUP) can be used. The weighting factors were calculated using marker effect derived Bayesian methods or iteratively based on single step marker effect. In multibreed genomic evaluation, there are several specific genetic structures into a genomic relationship matrix (G). The block wise genomic relationship matrix (BG) consist of several specific relationship blocks for each and pair breeds. BG can more accurately calculated relationships among animals than G matrix in multibreed genomic evaluations. The aim of this study is comparison G and BG and weighted BG (WBG) in weighted single step GWAS.

Materials and Methods: To conduct our study, we initially simulated two distinct populations, labeled as A and B, utilizing the QMsim software. The simulation involved the creation of two chromosomes, each spanning a length of two morgans. Within each chromosome, we simulated 2500 single nucleotide polymorphisms (SNPs). Subsequently, four traits were simulated, each possessing heritabilities of 0.05 and 0.3, along with varying numbers of quantitative trait loci (QTLs) set at 50 and 500. Following the simulation, we calculated the genetic value (G), the breeding value given by markers (BG), and the weighted breeding value given by markers (WBG) using SNP genotypes for all animals in the study. This comprehensive approach allowed us to evaluate and analyze the genetic and breeding values associated with the simulated traits across the populations. Genomic relationship matrices were used for single step GWAS (SSGWAS) analysis for each trait. 10 iterations was considered for single step SNP effect analyses. Moreover, the SNP effects were obtained by BayesB approach. BayesB effects was used for calculated weighting factors in WBG. Accuracies of methods and number of identified SNPs with explained genetic variance higher than 1% were reported.

Results and Discussion: using G and BG and WBG in SSGWAS led to identify 14, 16 and 21 SNP with higher than one percentage variance explained, respectively. Moreover, convergent accuracies of WssGBLUP using G and BG and WBG were 0.36, 0.39 and 0.43, respectively. WssGBLUP using WBG could be converged faster than using G and BG. Furthermore, accuracy of WssGBLUP using WBG was significantly more than

https://doi.org/10.22067/ijasr.2023.79619.1108

using G and BG. Multibreed GWAS is led to increase power of model because phenotypic information is severely increase. In multibreed GWAS, relationships among breeds usually are rare or zero but there are several locations among breeds that shared among them and should use those for genomic relationship calculation. In WBG and BG could be accurately calculate pair breeds genomic relationships using sharing pair breeds genomic locations. Principal component analyses showed that WBG was let to strongly increase genomic relationship among animals that is led to improved power of WssGWAS.

Conclusion: According to recent studies, multibreed genomic evaluation with the WssGBLUP can improve the accuracy of multibreed genomic evaluation, and the results of our study showed that for multibreed genomic evaluation and WssGWAS with the WssGBLUP , instead of the genomic relationship matrix (G), BG or WBG genomic relationship matrices are the better to use**.**

Keywords: Multibreed GWAS**,** Block wise genomic relationship matrix, WssGWAS, WssGBLUP, Explained genetic variance

نشریه پژوهشهای علوم دامی ایران

Homepage: http://ijasr.um.ac.ir

مقاله پژوهشی

جلد ۱۵، شماره ۴، زمستان ۱۴۰۲، ص. ۵۹۷-۵۸۵

مقایسه ماتریس‌های روابط خویشاوندی ژنومی متفاوت در پویش ژنومی وزن‌دهی شده چند جمعیتی

سید مصطفی مظلوم[1] و محمدمهدی شریعتی[2]*

تاریخ دریافت: ۱۴۰۱/۰۷/۲۴

تاریخ پذیرش: ۱۴۰۲/۰۲/۳۱

چکیده

هدف از پژوهش حاضر، انتخاب بهترین ماتریس روابط خویشاوندی ژنومی برای پویش ژنومی چندجمعیتی است. ساختار ژنتیکی جمعیت‌ها منحصر به فرد هستند، بنابراین برای تشکیل ماتریس روابط خویشاوندی ژنومی چند جمعیتی از ماتریس روابط خویشاوندی ژنومی بلوک‌بندی شده استفاده می‌شود. در این مطالعه، ابتدا دو جمعیت با ساختار ژنتیکی متفاوت شبیه‌سازی شدند. سپس ماتریس روابط خویشاوندی ژنومی معمولی (G)، ماتریس روابط خویشاوندی ژنومی بلوک‌بندی شده (BG) و ماتریس روابط خویشاوندی ژنومی بلوک‌بندی وزن‌دهی شده با واریانس ژنوتیپی برآورد شده با بیز B (WBG) برای حیوانات تشکیل شدند و برای مطالعات پویش ژنومی تک مرحله‌ای استفاده شدند. علاوه‌بر آن، با روش بیزB پویش ژنومی انجام شد و با پویش‌های ژنومی تک مرحله‌ای مقایسه شدند. نتایج پویش ژنومی با استفاده از ماتریس‌های روابط خویشاوندی G، BG و WBG نشان داد که به طور میانگین به‌ترتیب ۱۴، ۱۶، ۲۱ نشانگر ژنومی مرتبط با QTLهای صفات شناسایی شدند که واریانس ژنتیکی توجیه شده بالای یک درصد دارند. همچنین با استفاده از روش آماری بیز B تنها دو نشانگر ژنومی با واریانس بالای یک درصد شناسایی شدند. علاوه‌براین، میانگین صحت‌های پیش‌بینی ارزش اصلاحی ژنومی در لحظه هم‌گرایی با استفاده از ماتریس‌های G، BG و WBG به‌ترتیب ۰/۳۶، ۰/۳۹، ۰/۴۳ برآورد شدند. نتیجه گیری کلی نشان‌داد که استفاده از ماتریس‌های روابط خویشاوندی ژنومی BG و WBG می‌تواند باعث بهبود پویش ژنومی چندجمعیتی یا چندنژادی شود.

واژه‌های کلیدی: بیزB، پویش ژنومی چند جمعیتی، ماتریس روابط خویشاوندی ژنومی بلوکی، ماتریس روابط خویشاوندی ژنومی

مقدمه

پویش گسترده ژنومی (GWAS) با استفاده از اطلاعات فنوتیپی صفات و ژنوتیپ نشانگرهای ژنومی (SNPs)، واریانت‌های ژنومی مرتبط با جایگاه‌های صفات کمّی (QTL) را شناسایی می‌کند. با افزایش دسترسی به توالی‌یابی ژنومی شناسایی واریانت‌ها و ژن‌های مرتبط با صفات اقتصادی در حیوانات اهلی نسبت به گذشته آسان‌تر شده است. علاوه‌براین، وجود عدم تعادل لینکاژی (LD) تقریباً گسترده بین نشانگرهای ژنومی و جایگاه‌های صفات کمّی و انتقال بلوک‌های هاپلوتایپی کامل طی نسل‌ها باعث شده است که با سهولت بیشتر نشانگرهای ژنتیکی مرتبط با صفات اقتصادی در حیوانات اهلی شناسایی شوند (Liu et al., 2023; van den Berg et al., 2016).

روش‌های آماری رگرسیون خطی و بیزی معمولاً برای GWAS در حیوانات و گیاهان و مطالعات بیماری‌های انسانی استفاده می‌شوند (Meuwissen *et al*., 2009; Reales and Wallace, 2023). هدف اصلی GWAS شناسایی جایگاه‌های صفت کمّی (QTL)با تأثیر عمده است. اخیراً با معرفی روش آماری پویش ژنومی تک مرحله‌ای (ssGWAS) برای برآورد واریانس ژنتیکی نشانگرهای ژنومی از پیش‌بینی ارزش اصلاحی به‌روش آماری پیش‌بینی ناارِیب خطی تک مرحله‌ای (ssGBLUP) استفاده می‌شود (Wang *et al*., 2012; Wijesena *et al*., 2023). پیش‌بینی ارزش اصلاحی ژنومی و واریانس ژنتیکی پلی‌مورفیسم تک نوکلئوتیدی (SNP) را می‌توان به‌راحتی از هر دو روش GBLUP و بیزی به‌دست آورد (Goddard

۱- دانشجوی دکتری ژنتیک و اصلاح نژاد دام، گروه علوم دامی، دانشکده کشاورزی، دانشگاه فردوسی مشهد، ایران.

۲- استادیار، گروه علوم دامی، دانشکده کشاورزی، دانشگاه فردوسی مشهد، ایران.

(*- نویسنده مسئول: (Email: mm.shariati@um.ac.ir

https://doi.org/10.22067/ijasr.2023.79619.1108

and Hayes, 2009). GBLUP یک توزیع نرمال را برای اثرات SNP فرض می‌کند و اثر آن‌ها را از روی فنوتیپ‌ها و ماتریس روابط خویشاوندی ژنومی (G) محاسبه می‌کند. هنگامی که فقط تعداد اندکی از افراد یک جمعیت تعیین ژنوتیپ می‌شوند، GBLUP تک مرحله‌ای (ssGBLUP) روش مناسبی برای پیش‌بینی ارزش اصلاحی ژنومی حیوانات است (Aguilar *et al*., 2010). اطلاعات اضافی در مورد حیوانات فاقد ژنوتیپ، صحت پیش‌بینی ارزش اصلاحی افزایش می‌دهد (Legarra *et al*., 2015).

روش‌های بیزی نظیر بیزA و بیزB روش‌هـای غیرخطـی هسـتند که برای برآورد اثر جایگزینی آللی هـر یـک از نشـانگرهای ژنـومی (SNPs)، توزیع پیشین واریانس اثرات ژنتیکـی در نظـر مـی‌گیرنـد و سپس با توزیع پسین واریانس اثرات ژنتیکی و از طریق زنجیره نمونه-گیری مونتوکارلو (MCMC) اثرات نشانگرهای ژنومی برآورد می‌شود (Meuwissen et al., 2009). امّـا روش‌هـای آمـاری GBLUP و ssGBLUP برای تمامی نشانگرهای ژنومی توزیع و وزن یکسـانی را در نظر می‌گیرند (Vanraden, 2008). توزیع اثرات ژن‌ها در صـفات اقتصادی یکسان نیست، بلکه در بسیاری از صـفات اقتصـادی تعـداد اندکی جایگاه صفت کمّی (QTL) عمده اثر هستند و دیگـر جایگـاه-های صفت کمّی اثـرات جزئـی دارنـد (Cesarani *et al*., 2022). در روش‌های بیزی با وزن‌دهی مخصـوص بـه‌عنـوان پـیش فـرض بـه نشانگرهای ژنومی، سعی بر این است که نشانگرهایی که با جایگـاه-های صفت کمّی (QTLs) عمده اثر پیوستگی دارند و اثـرات بزرگـی دارند، شناسایی شـوند (Su *et al*., 2014; Zhang *et al*., 2010). مطالعات شبیه‌سازی نشان داده است، زمـانی کـه صـفت دارای تعـداد اندکی QTL با اثرات بزرگ است، روش‌های آماری بیزی نسـبت بـه GBLUP کارایی بالاتری برای پیش‌بینـی ارزش اصـلاحی ژنـومی و پویش ژنومی دارنـد (Lund et al., 2009 ; Gao et al., 2012). در مطالعات واقعی در گاوشیری نتایج نشان داده است، بـه‌دلیـل فراوانـی بالای QTLها در صفات اقتصادی روش‌های بیزی صحت پیش‌بینـی ارزش اصـلاحی ژنـومی کمتـری نسـبت بـه روش‌هـای GBLUP و ssGBLUP دارند (Su et al., 2010 ;Cole et al., 2009).

یکی از روش‌های گنجاندن واریانس اختصاصی برای نشانگرهای ژنومی در روش‌آماری GBLUP، وزن‌دهـی نشـانگرهای ژنـومی بـه هنگام تشکیل ماتریس روابط خویشاوندی ژنومی (G) است (Zhang et al., 2010). در این روش‌آماری (WGBLUP)، پیش‌بینـی ارزش اصلاحی تک مرحله‌ای با توزیع واریانس ژنتیکی نشـانگرهای ژنـومی مشابه روش‌های بیزی انجام می‌شود و هر نشانگر ژنـومی بـر اسـاس میـزان پیوسـتگی و ارتبـاطی کـه بـا QTLهـای صـفت دارد، وزن مخصوص به خود می‌گیرد و اصطلاحاً مـاتریس روابـط خویشـاوندی وزن‌دهی اختصاصی (GW) می‌شود. در بسیاری مطالعات نشـان داده شده است که وزن‌دهی ماتریس روابط خویشاوندی ژنـومی مـی‌توانـد منجر به بهبود پیش‌بینی ارزش اصلاحی ژنومی شـود (Tiezzi and Maltecca, 2015 ;Zhang et al., 2014 ;Su et al .,2014).

چندین روش برای ساخت ضرایب وزن‌دهی نشـانگرهای ژنـومی معرفی شده است. به‌عنوان مثال، روش‌های اسـتفاده از نتـایج آزمـون رگرسـیونی پـویش ژنـومی ((-log(p-value)) و اسـتفاده از واریـانس ژنتیکی نشانگرهای ژنـومی بـرآورد شـده بـا اثـرات جـایگزینی آللـی نشانگرهای ژنـومی ($2pq\alpha^2$) بـرای وزن‌دهـی نشـانگرهای ژنـومی استفاده مـی‌شـوند(Gualdrón Duarte *et al*., Su et al., 2014; 2020). اخیراً، مطالعات وزن‌دهی ماتریس روابط خویشـاوندی ژنـومی نشان داده است که وزن‌دهی نشانگرهای ژنومی با استفاده از واریانس ژنتیکی بـرآورد شـده بـا اثـرات نشـانگرهای ژنـومی بـرآورد شـده از پیش‌بینی ارزش اصلاحی ژنومی GBLUP و ssGBLUP مـی‌توانـد باعث بهبود پویش ژنومی شود (Wang *et al*., 2014).

ونگ و همکاران (Wang *et al*., 2014) با شبیه‌سازی داده‌هـای ژنومی نشان دادند که با روشی چند مرحله‌ای وزن‌دهـی نشـانگرهای ژنومی با اثرات برآورد با استفاده از ارزش‌های اصلاحی پیش‌بینی شده بـا روش WssGBLUP و بـرآورد اثـرات نشـانگری بـا اسـتفاده از اطلاعات آخرین تکرار وزن‌دهی باعث می‌شود که محدوده QTLها را با دقت بیشتری نسبت به روش‌های بیزی نظیر بیز B شناسـایی کـرد (*Zhang et al., 2016*) در روش پویش ژنومی وزن‌دهی شـده تـک مرحله‌ای (WssGWAS) به‌دلیل اینکه همزمان از اطلاعات حیوانات ژنوتیپ شده و فاقد ژنوتیپ برای پـیش‌بینـی ارزش اصـلاحی ژنـومی استفاده می‌شود و صحت پیش‌بینی ارزش اصـلاحی ژنـومی افـزایش قابل ملاحظه‌ای دارد، می‌تواند بـا دقـت و تـوان بـالاتری نسـبت بـه روش‌های بیزی جایگاه‌های صفت کمّی را شناسایی نماید (Wang et al., 2014). مطالعات پویش ژنومی و انتخاب ژنومی اخیر نشـان داده است که افزایش اندازه جمعیت می‌تواند به تـوان روش آمـاری بـرای شناسایی QTLهای صفت و پیش‌بینی ارزش اصلاحی ژنومی کمـک کند. بنابراین، مطالعات ژنومی چند نژادی یا چند جمعیتی بسـیار حـائز اهمیـت شـده اسـت (Erbe et al., 2012; Raven et al., 2014).

وارونز و همکاران (Veroneze *et al*., 2016)، به روش‌های معمولی و بلوک‌بندی شده ماتریس‌های روابـط خویشـاوندی ژنـومی تشـکیل دادنـد و بـه‌روش GBLUP بـرای حیوانـات دو نـژاد خـوک ارزش اصلاحی ژنومی پیش‌بینی کردند. نتایج آن‌ها نشان داد که اسـتفاده از ماتریس روابط خویشاوندی ژنـومی بلـوک‌بنـدی شـده (BG) باعـث افزایش صحت پیش‌بینی ارزش اصلاحی ژنومی قابل ملاحظه‌ای شده است، همچنین در این مطالعـه مـاتریس روابـط خویشـاوندی ژنـومی بلوک‌بندی شده وزن‌دهی (WBG) شد و نسبت به غیر وزن‌دهی شده (BG) آن صحت پیش‌بینی ارزش اصلاحی ژنومی بالاتری نشان داد. در مطالعه‌ای دیگر، آلوارنگا و همکـاران (Alvarenga et al., 2020) با شبیه‌سازی، جمعیت‌های آمیخته و خالص پدری و مـادری در قالـب

یک ماتریس روابط خویشاوندی ژنومی (G) وزن‌دهی کردند. نتایج آن‌ها نشان داد که استفاده از ماتریس روابط خویشاوندی ژنومی وزن-دهی شده معمولی (GW) نمی‌تواند صحت پیش‌بینی ارزش اصلاحی ژنومی چند جمعیتی را افزایش دهد. با توجه به مطالعات اخیر، استفاده از چند جمعیت به‌جای یک جمعیت به‌عنوان جمعیت مرجع با استفاده از ماتریس روابط خویشاوندی ژنومی بلوک‌بندی شده (BG) یا وزن‌دهی شده آن (WBG) می‌تواند پیش‌بینی ارزش اصلاحی ژنومی را بهبود ببخشد (Kemper et al., 2015; Veroneze et al., 2016). هدف این مطالعه، انتخاب بهترین ماتریس روابط خویشاوندی ژنومی برای برآورد غیرخطی واریانس ژنتیکی برآورد شده با اثرات جایگزینی آللی نشانگرهای ژنومی به‌روش آماری WssGWAS برای پویش ژنومی چند جمعیتی (استفاده از اطلاعات فنوتیپی و ژنوتیپی حیوانات ژنوتیپ شده و فاقد ژنوتیپ بیش از یک جمعیت) است. بدین ترتیب ابتدا دو جمعیت با ساختار ژنتیکی کاملاً مجزا شبیه‌سازی شدند. سپس با استفاده از سه ماتریس روابط خویشاوندی ژنومی (G)، بلوک‌بندی شده (BG) و بلوک‌بندی شده وزن‌دهی شده (WBG) به روش WssGWAS با استفاده از چرخه وزن‌دهی ۱۰ مرحله‌ای، واریانس ژنتیکی برآورد شده با اثرات جایگزینی آللی نشانگرهای ژنومی برآورد شدند و پویش ژنومی چند جمعیتی انجام شد. سپس صحت‌های پیش‌بینی ارزش اصلاحی ژنومی تمامی مرحله‌ها برآورد و گزارش شدند (Zhang *et al*., 2016).

مواد و روش‌ها

شبیه‌سازی ژنوم و جمعیت

در این مطالعه، چهار صفت کمّی با وراثت‌پذیری‌های ۰/۰۵ و ۰/۳ و تراکم‌های جایگاه صفت کمّی (QTLs) ۵۰ و ۵۰۰ با واریانس فنوتیپی یک روی ژنومی متشکل از دو کروموزوم، هر یک به‌طول ۱۰۰ سانتی مورگان شبیه‌سازی شدند. روی هر یک از کروموزوم‌ها ۵۰۰۰ نشانگر ژنومی (SNPs) توزیع شدند. موقعیت جایگاه‌های صفت کمّی و نشانگرهای ژنومی تصادفی انتخاب شدند. اثرات جایگزینی آللی برای QTLها از یک توزیع گاما با پارامتر شکلی ۰/۴ و نرخ ۱/۶۶ نمونه‌گیری شدند و برای هر QTL تعداد دو، سه یا چهار آلل به تصادف شبیه‌سازی شد، امّا نشانگر ژنومی دو آللی (چند شکلی تک نوکلئوتیدی) شبیه‌سازی شدند. نرخ جهش برگشت‌پذیر برای جایگاه‌های صفت کمّی QTLs و نشانگرهای ژنومی SNPs 10^{-6} در نظر گرفته شد.

برای ایجاد عدم تعادل لینکاژی (LD) در فاصله ژنتیکی کوتاه به روش تعادل جهش – رانش ژنتیکی، ابتدا یک جمعیت پایه به اندازه ۱۰۰۰ حیوان به‌مدت ۵۰۰ نسل آمیزش تصادفی داده شدند و سپس اندازه جمعیت به ۵۰۰۰ حیوان به‌مدت ۵۰۰ نسل گسترش داده شد. در نتیجه این فرآیند ساخت جمعیت منجر به شبیه‌سازی جمعیت تاریخی شد. برای ایجاد LD در فاصله ژنتیکی بلند و ایجاد ساختار ژنتیکی جمعیت ثانویه، ۱۰۰۰ حیوان نر و ۴۰۰۰ حیوان ماده نسل آخر جمعیت تاریخی به‌مدت ۵۰ نسل با یکدیگر آمیزش تصادفی داده شدند. نرخ جایگزینی والدین در جمعیت ثانویه ۰/۵ در نظر گرفته شد و طرح انتخاب و حذف والدین به‌ترتیب تصادفی و سن بالا در نظر گرفته شد. سپس برای ایجاد دو جمعیت مشتق شده، برای جمعیت اول (A) از بین حیوانات نسل آخر جمعیت ثانویه ۲۰ حیوان نر و ۱۰۰ حیوان ماده (نسل پایه) انتخاب شدند و برای جمعیت دوم (B) از بین حیوانات نسل ۲۰ جمعیت ثانویه ۳۰ حیوان نر و ۱۲۰ حیوان ماده (نسل پایه) انتخاب شدند. هر دو جمعیت به‌مدت ۵۰ نسل آمیزش تصادفی داده شدند. نرخ جایگزینی والد نر در جمعیت‌های اول و دوم به‌ترتیب ۸۰ و ۵۰ درصد انتخاب شدند و نرخ جایگزینی والد ماده در جمعیت اول و دوم به‌ترتیب ۳۰ و ۱۰ درصد انتخاب شدند. نرخ رشد تعداد والد نر جمعیت اول و دوم ۱۰ درصد و نرخ رشد تعداد والد ماده جمعیت‌های اول و دوم به‌ترتیب ۱۰ درصد و پنج درصد جمعیت پایه انتخاب شدند. انتخاب حیوانات برای جایگزینی والدین در جمعیت‌های اول و دوم به ترتیب ارزش اصلاحی بالا و ارزش فنوتیپی بالا در نظر گرفته شد و حیوانات مسن از جمعیت‌ها حذف شدند و نرخ نتاج متولد شده به‌ازاء هر حیوان ماده یک فرزند انتخاب شد (مطابق با گله‌های گاوهای نژاد شیری). میزان LD در فواصل کوتاه و بلند در نسل آخر جمعیت‌های اول و دوم به‌ترتیب ۰/۳۵ و ۰/۲۸ و ۰/۱۱ و ۰/۰۸ برآورد شدند. این مقادیر تقریباً مشابه مقادیر LD گزارش شده در حیوانات اهلی بودند (De Roos et al., 2008).

ژنوتیپ نشانگرهای ژنومی حیوانات نسل آخر جمعیت‌های اول و دوم ذخیره شدند. علاوه‌براین، شجره و ارزش‌های فنوتیپی سه نسل آخر جمعیت اول و پنج نسل آخر جمعیت دوم ذخیره شدند. در مجموع، اطلاعات شجره و ارزش فنوتیپی ۴۷۹۸ حیوان ذخیره شدند که ۱۲۹۸ حیوان دارای ژنوتیپ بودند. برای مقایسه روش‌های آماری و اعتبارسنجی بهبود پویش ژنومی شبیه‌سازی با ۱۰ تکرار انجام شد. ژنوتیپ‌های نشانگرهای ژنومی مطابق کمترین فراوانی آللی (MAF) بیشتر از یک درصد و آزمون تعادل هاردی- وینبرگ با (P-value>10^{-6}) با استفاده از نرم‌افزار PLINK (Purcell *et al*., 2007) کنترل کیفیت شدند.

مدل آماری

مدل آماری مختلط برای پیش‌بینی ارزش اصلاحی ژنومی برای پویش ژنومی با حیوانات ژنوتیپ شده و غیر ژنوتیپ شده، مدل آماری WssGBLUP است (Legarra et al., 2014)، که با استفاده از ماتریس روابط خویشاوندی ترکیبی شجره و ژنومی ارزش‌های

اصلاحی حیوانات و اثرات جایگزینی آللی نشانگرهای ژنومی را برآورد می‌کند، به شکل زیر است:

معادله (۱)
$$y = Xb + Zg + e$$

در معادله ۱، y: بردار ارزش‌های فنوتیپی، ماتریس‌های X و Z: به‌ترتیب ارزش‌های فنوتیپی به اثرات ثابت نژاد و حیوان‌ها مرتبط می‌کنند. b و g: به‌ترتیب بردارهای اثرات ثابت و ارزش‌های اصلاحی حیوانات با توزیع $g \sim N(0, H\sigma_g^2)$ هستند. e: بردار اثرات باقیمانده با توزیع $e \sim N(0, I\sigma_e^2)$ است. H و I: به‌ترتیب ماتریس روابط خویشاوندی ترکیبی (نحوه تشکیل آن در ادامه توضیح داده خواهد شد) و ماتریس قطری واحد هستند. همچنین، σ_g^2 و σ_e^2: به‌ترتیب واریانس‌های ژنتیک افزایشی و باقی‌مانده هستند. مؤلفه‌های واریانس، ارزش‌های اصلاحی حیوانات و اثرات جایگزینی آللی نشانگرهای ژنومی با استفاده از نرم‌افزارهای BLUPF90 و postGSF90 انجام شدند (Lourenco et al., 2020).

تشکیل ماتریس‌های روابط خویشاوندی ترکیبی شجره و ژنومی

ماتریس H یک ماتریس روابط خویشاوندی ترکیبی است، که در آن رابطه خویشاوندی بین افراد دارای ژنوتیپ و بدون ژنوتیپ برآورد می‌شود. در واقع، یک ماتریس واحد برای تشکیل روابط خویشاوندی کل افراد یک جمعیت با شجره و ژنوتیپ نشانگرهای ژنومی است. ماتریس H^{-1} در معادله ۲ آمده است (Legarra *et al.*, 2014).

معادله (۲)
$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & Gw^{-1} - A_{22}^{-1} \end{bmatrix}$$

که در آن، A^{-1}، Gw^{-1} و A_{22}^{-1} به‌ترتیب معکوس ماتریس‌های روابط خویشاوندی شجره کل حیوانات، روابط خویشاوندی ژنومی, حیوانات دارای ژنوتیپ (در ادامه نحوه تشکیل آن توضیح داده خواهد شد) و روابط خویشاوندی شجره حیوانات دارای ژنوتیپ هستند. در این مطالعه، هدف انتخاب بهترین ماتریس روابط خویشاوندی ژنومی برای تشکیل ماتریس روابط خویشاوندی ترکیبی است که بتواند صحت پیش‌بینی ارزش‌های اصلاحی و دقت پویش ژنومی چند جمعیتی را بهبود بخشد.

با توجه به اینکه ساختار ژنتیکی جمعیت‌ها با یکدیگر متفاوت هستند و ساختار ژنومی آن‌ها نظیر LD و فراوانی آللی نشانگرهای ژنومی کاملاً متفاوت هستند. در نتیجه، باید روابط خویشاوندی ژنومی حیوانات درون هر جمعیت جدا از روابط خویشاوندی بین حیوانات دو جمعیت برآورد شود (Wientjes *et al.*, 2017). اصطلاحاً در هنگام تشکیل ماتریس روابط خویشاوندی ژنومی، روابط خویشاوندی حیوانات هر جمعیت و روابطه خویشاوندی بین حیوانات دو جمعیت در بلوک‌های خویشاوندی کاملاً مجزا برآورد می‌شوند (Zhou *et al.*, 2014). بنابراین، در این مطالعه به دو روش معمولی و بلوک‌بندی شده، ماتریس‌های روابط خویشاوندی ژنومی تشکیل شدند و معکوس آن‌ها برای تشکیل ماتریس H^{-1} استفاده شدند. روش معمولی (G) به صورت معادله ۳ (Vanraden, 2008) و روش بلوک‌بندی شده (BG) به‌صورت معادله ۴ (Wientjes et al., 2017) تشکیل شدند.

معادله (۳)
$$G = \frac{MD_tM'}{2\sum 2p_{AB}(1-p_{AB})}, G^{-1} = (0.95G + 0.05A_{22})^{-1}$$

معادله (۴)
$$BG = \begin{bmatrix} \frac{M_A D_t M_A'}{2\sum p_A(1-p_A)} & \frac{M_A D_t M_B'}{2\sum \sqrt{p_A(1-p_A)p_B(1-p_B)}} \\ \frac{M_B D_t M_A'}{2\sum \sqrt{p_B(1-p_B)p_A(1-p_A)}} & \frac{M_B D_t M_B'}{2\sum p_B(1-p_B)} \end{bmatrix}$$

$$BG^{-1} = (0.95BG + 0.05A_{22})^{-1}$$

در معادله ۳، M و p_{AB} به‌ترتیب ماتریس ژنوتیپی مرکزی شده و فراوانی آللی دو جمعیت A و B هستند. در معادله ۴: M_A و M_B و p_A و p_B به‌ترتیب ماتریس ژنوتیپی مرکزی شده و فراوانی آللی اختصاصی هر یک از جمعیت‌های A و B هستند. در هر دو معادله، D_t: ماتریس قطری ضرایب وزن‌دهی در تکرار t ام است. در روش آماری WssGWAS ابتدا در تکرار اول D=I است و در مراحل بعدی با اثرات نشانگرهای ژنومی برآورد شده برای هر نشانگر ژنومی ضریب وزن‌دهی اختصاصی برآورد می‌شود و این تکرارها تا زمانی که صحت‌های پیش‌بینی ارزش اصلاحی ژنومی به همگرایی برسند، یعنی زمانی که دیگر با وزن‌دهی صحت تغییر قابل ملاحظه‌ای نکند، ادامه دارد (Zhang *et al.*, 2016). سپس اثرات نشانگرهای ژنومی برآورد شده در تکرار آخر برای پویش ژنومی استفاده می‌شوند. در این مطالعه، با استفاده از ماتریس‌های روابط خویشاوندی ژنومی به روش‌های معادله‌های ۵ (با استفاده از ماتریس G) (Wang *et al.*, 2012) و ۶ (با استفاده از ماتریس BG) (Wientjes *et al.*, 2017; Wang *et al.*, 2012) برای نشانگرهای ژنومی اثرات جایگزینی آللی برآورد شدند. سپس برای تکرار بعد به‌روش معادله ۷ برای نشانگرهای ژنومی ضرایب وزن‌دهی برآورد شدند. علاوه‌براین، از

ماتریس روابط خویشاوندی ژنومی بلوک‌بندی وزن‌دهی شده با ضرایب وزن‌دهی برآورده شده با اثرات نشانگری مدل آماری بیز B (WBG) (Su *et al*., 2014; Veroneze et al., 2016) نیز برای پویش ژنومی و مقایسه با ماتریس‌های روابط خویشاوندی ژنومی غیر وزن‌دهی استفاده شد. همچنین برای بررسی کارآمدی پویش ژنومی چند جمعیتی با استفاده از روش آماری WssGWAS، با استفاده از مدل آماری بیز B با ضریب احتمال تأثیر $\pi = 0.95$ با استفاده از پکیج BGLR (Pérez and De Los Campos, 2014) و الگوریتم نمونه‌گیری گیبس پویش ژنومی چند جمعیتی انجام شد (Zhang *et al*., 2016).

معادله (۵)

$$\hat{a_t} = \alpha D_t M' G^{-1} \hat{g}_t \,, \alpha = \frac{1}{2\sum p_{AB}(1-p_{AB})}$$

معادله (۶)

$$\hat{a_t} = \beta D_t M' BG^{-1} \hat{g}_t \,, \beta = \frac{1}{2\sum \sqrt{p_A(1-p_A)p_B(1-p_B)}}$$

معادله (۷)

$$D_{t+1} = \mathrm{Diag}\left(\frac{nV_{i_t}}{\sum V_{i_t}}\right), Vi_{BG} = 2\sqrt{p_A(1-p_A)p_B(1-p_B)}\hat{a}^2, Vi_G = 2p_{AB}(1-p_{AB})\hat{a}^2$$

در معادله ۵ تا ۷، $\hat{a}$: بردار اثرات جایگزینی آللی نشانگرهای ژنومی برای هر تکرار t ام، M: ماتریس مرکزی شده ژنوتیپ‌های نشانگرهای ژنومی، α و β: ضرایب تصحیح هستند. Vi_G و Vi_{BG}: به ترتیب واریانس‌های ژنوتیپی برآورد شده با اثرات نشانگرهای ژنومی به‌دست آمده از ماتریس‌های روابط خویشاوندی ژنومی G و BG هستند. برای محاسبه میزان واریانس ژنتیکی توجیه شده توسط هریک از نشانگرهای ژنومی از معادله ۸ استفاده شد. برای مقایسه ماتریس‌های روابط خویشاوندی ژنومی، صحت پیش‌بینی ارزش اصلاحی ژنومی و تعداد نشانگرهای ژنومی و مجموع واریانس توجیه شده توسط نشانگرهای ژنومی که حاوی بیش از یک درصد واریانس ژنتیکی کل هستند، برآورد شدند.

معادله (۸)

$$gVar\%_i = \frac{2p_{AB\,i}(1-p_{AB\,i})\hat{a}^2{}_i}{\sigma_a^2} * 100$$

در معادله ۸، $gVar\%$ و p_{AB}: به‌ترتیب درصد واریانس ژنتیکی توجیه شده و فراوانی آللی نشانگر ژنومی i ام هستند. σ_a^2: واریانس ژنتیکی کل است (Wang *et al*., 2012).

نتایج و بحث

آماره‌های توصیفی ارزش‌های فنوتیپی صفات شبیه‌سازی شده برای هر یک از جمعیت‌ها در جدول ۱ نشان داده شده است. نتایج نشان می‌دهد، نسبت به جمعیت پایه با افزایش نسل تعداد QTLهای مؤثر بر روی صفات کاهش یافته است که نشان می‌دهد، دو جمعیت تحت انتخاب بوده‌اند. همچنین، این تعداد QTLها، جایگاه‌هایی هستند که بین دو جمعیت مشترک می‌باشند. همچنین به‌دلیل اینکه نحوه و شدت انتخاب بین جمعیت‌ها متفاوت است، میزان وراثت‌پذیری صفت در دو جمعیت متفاوت مشاهده شدند. علاوه‌بر این‌ها، نتایج نشان داد که ارزش‌های فنوتیپی صفات شبیه‌سازی شده در دو جمعیت A و B کاملاً نسبت به یکدیگر ناهمگن بوده و مجزا می‌باشند. در شبیه‌سازی صفات سعی شد که وراثت‌پذیری‌ها منطبق با صفات اقتصادی در گله‌های گاو شیری باشند (Pedrosa *et al*., 2023).

جدول ۱- آمار توصیفی ارزش فنوتیپی صفات شبیه‌سازی شده

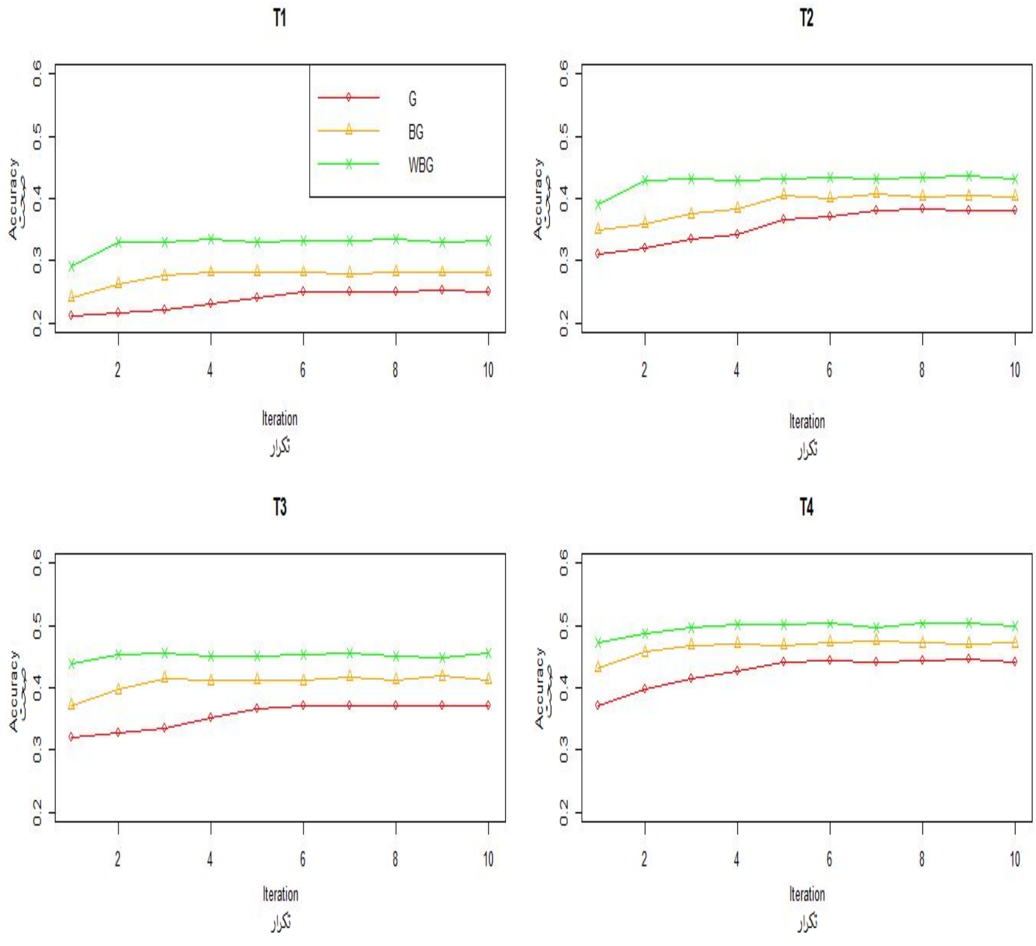
Table 1- Descriptive statistics of phenotypic values of simulated traits

صفات Traits	تعداد جایگاه صفت کمّی QTL Number of QTLs	جمعیت Population	وراثت‌پذیری h2	حداکثر Maximum	حداقل Minimum	میانگین Mean	واریانس Variance
T1	47	A	0.015±0.001	3.6±0.05	-2.31±0.01	0.91±0.01	0.97±0.01
		B	0.032±0.002	3.7±0.04	-2.19±0.02	0.79±0.02	0.88±0.02
T2	482	A	0.013±0.001	4.5±0.03	-1.24±0.01	1.64±0.03	0.95±0.01
		B	0.029±0.002	4.6±0.01	-1.89±0.01	0.87±0.01	0.98±0.03
T3	43	A	0.18±0.01	5.4±0.02	-0.03±0.01	2.68±0.02	0.71±0.01
		B	0.23±0.01	5±0.03	-0.28±0.01	2.28±0.01	0.73±0.01
T4	478	A	0.15±0.02	7.2±0.07	2.39±0.01	4.95±0.01	0.75±0.02
		B	0.22±0.01	7±0.08	0.86±0.01	4.15±0.01	0.88±0.01

جدول ۲- صحت‌های پیش‌بینی ارزش اصلاحی ژنومی

Table 2- Accuracies of genomic prediction

صفات Traits	ماتریس روابط خویشاوندی ژنومی G	ماتریس روابط خویشاوندی ژنومی بلوک‌بندی شده Blok wise G	ماتریس روابط خویشاوندی ژنومی وزن‌دهی شده بلوک‌بندی Weighted blok wise G	بیز B BayesB
T1	0.25±0.01	0.28±0.01	0.33±0.01	0.24±0.01
T2	0.38±0.03	0.4±0.03	0.43±0.03	0.32±0.02
T3	0.37±0.01	0.41±0.01	0.45±0.01	0.38±0.01
T4	0.44±0.02	0.47±0.02	0.5±0.02	0.4±0.02



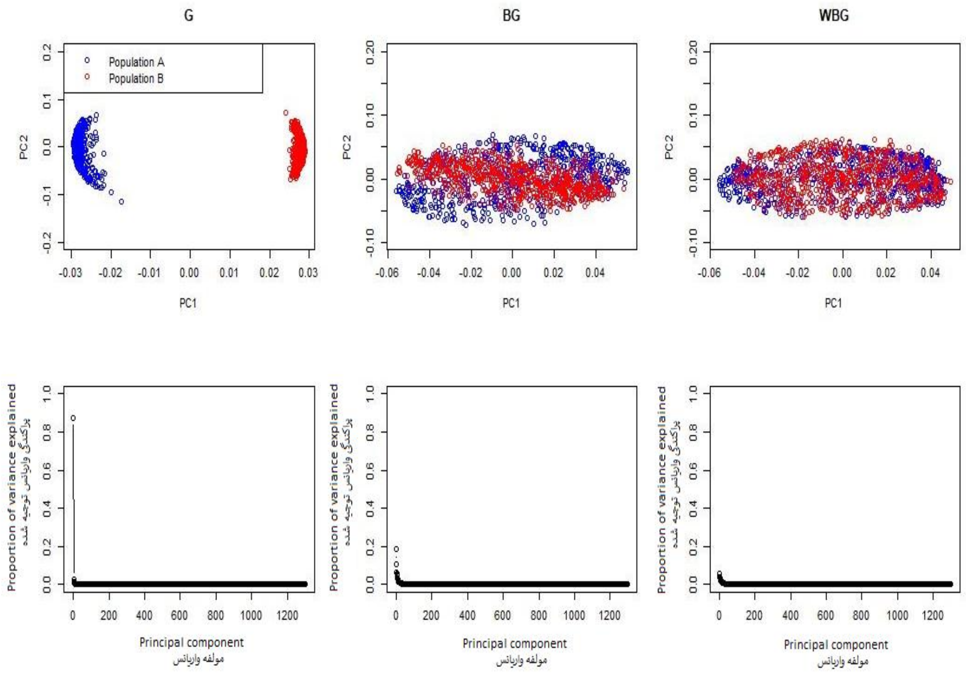
شکل ۱- نمودارهای تغییرات صحت در تکرارهای مختلف پویش ژنومی چند جمعیتی برای ماتریس‌های ژنومی متفاوت در صفات T1 تا T4

Figure 1- Plots of iterations of accuracies from T1 to T4 traits

نتایج برآورد صحت در زمان همگرایی تکرارها در مطالعات پویش ژنومی تک مرحله‌ای نشان داد، که استفاده از ماتریس روابط خویشاوندی ژنومی بلوک‌بندی شده (BG) و بلوک‌بندی وزن‌دهی شده (WBG) می‌تواند صحت پیش‌بینی ارزش اصلاحی ژنومی به روش WssGBLUP چند جمعیتی را بهبود بخشد (جدول ۲). علاوه‌براین نتایج نشان داد که روش آماری بیز B چند جمعیتی به دلیل متفاوت بودن فاز LD در جمعیت‌های مختلف و همچنین به‌دلیل اطلاعات فنوتیپی یا مشاهدات کمتر از مدل‌های ssGBLUP و WssGBLUP نسبت به مدل‌های گفته شده کارایی کمتری دارد و صحت پیش‌بینی ارزش اصلاحی این روش کمتر از روش‌های گفته شده است (Zhang et al., 2016) همچنین نتایج نشان داد، با افزایش تراکم QTL صحت پیش‌بینی ارزش اصلاحی افزایش یافت، که این امر به‌دلیل آن است که ماتریس روابط خویشاوندی ژنومی در حالت چند ژنی کارایی بیشتری نسبت به حالت تک ژنی صفات دارد

(Goddard and Hayes, 2009; Vanraden, 2008; Wientjes *et al*., 2015). استفاده از ماتریس روابط خویشاوندی ژنومی بلوک‌بندی وزن‌دهی شده (WBG) برای پیش‌بینی ارزش اصلاحی به‌روش WssGBLUP چند جمعیتی نسبت به ماتریس‌های BG و G به‌صورت میانگین باعث به‌ترتیب ۱۸ و ۳۲ درصد افزایش صحت پیش‌بینی ارزش اصلاحی شده است. همچنین در هنگام برآورد اثرات نشانگرهای ژنومی به‌روش WssGWAS نسبت به سایر ماتریس‌های روابط خویشاوندی ژنومی در زمان کمتری به همگرایی صحت پیش‌بینی ارزش اصلاحی ژنومی می‌رسد (شکل۱). ژنگ و همکاران (Zhang *et al*., 2016) نشان دادند که استفاده از ماتریس روابط خویشاوندی ژنومی GW می‌تواند سرعت و عملکرد WssGWAS تک جمعیتی را بهبود بخشد. علاوه‌بر آن، اخیراً مطالعات نشان دادند که استفاده از ماتریس WBG می‌تواند صحت پیش‌بینی ارزش اصلاحی ژنومی با روش GBLUP چند جمعیتی را نسبت به ماتریس G افزایش دهد (Veroneze et al., 2016). در مطالعه‌ای دیگر نشان داده شد که استفاده از ماتریس روابط خویشاوندی ژنومی GW در روش ssGBLUP چند جمعیتی در جوامع آمیخته نمی‌تواند منجر به بهبود صحت پیش‌بینی ارزش اصلاحی شود (Alvarenga et al., 2020).

نتایج رسم نمودار PCA ماتریس‌های روابط خویشاوندی در تکراری که صحت پیش‌بینی ارزش اصلاحی به همگرایی می‌رسد، نشان داد که ماتریس WBG نسبت به ماتریس‌های BG و G توانسته است، فاصله ژنتیکی بین جمعیت را کاهش دهد و این امر باعث افزایش توان روش پویش ژنومی چند جمعیتی و افزایش صحت پیش‌بینی ارزش اصلاحی ژنومی شده است (Wu et al., 2015). به طوری‌که در PC1 میزان واریانس برای ماتریس‌های WBG، BG و G به‌ترتیب ۰/۰۷، ۰/۲ و ۰/۸ برآورد شد (شکل ۲). این نتایج نشان داد که ماتریس WBG توانسته است، نسبت به سایرین بیشتر از اطلاعات نواحی ژنومی مشترک بین حیوانات دو جمعیت A و B برای همگن‌سازی خویشاوندی ژنومی چند جمعیتی استفاده نماید (Veroneze et al., 2016). دیتویلر و همکاران (Daetwyler *et al*., 2012) نشان دادند که نزدیکی و دوری رابطه خویشاوندی ژنومی بین حیوانات مرجع و آزمون بر روی صحت و اریبی پیش‌بینی ارزش اصلاحی مؤثر هستند و هر چقدر روابط خویشاوندی ژنومی بین حیوانات بیشتر باشد، صحت پیش‌بینی ارزش اصلاحی بیشتر خواهد بود.



شکل ۲– نمودارهای تجزیه و تحلیل روابط خویشاوندی ژنومی برای ماتریس‌های روابط خویشاوندی ژنومی مختلف
Figure 2- PCA plots of different genomic relationship matrices

مهم‌ترین عامل افزایش توان روش‌های آماری پویش ژنومی اندازه جمعیت مرجع است (Raven et al., 2014). حیوانات اقتصادی نظیر گاو شیری تحت انتخاب شدید هستند. همچنین در نژادهای گاو شیری معمولاً جهت و نحوه انتخاب‌ها نزدیک به یکدیگر است ، که باعث می‌شود نواحی ژنومی تحت انتخاب مشترک در آن‌ها مشاهده شود. بنابراین، به لحاظ فاز LD نواحی مشترک زیادی بین آن‌ها وجود

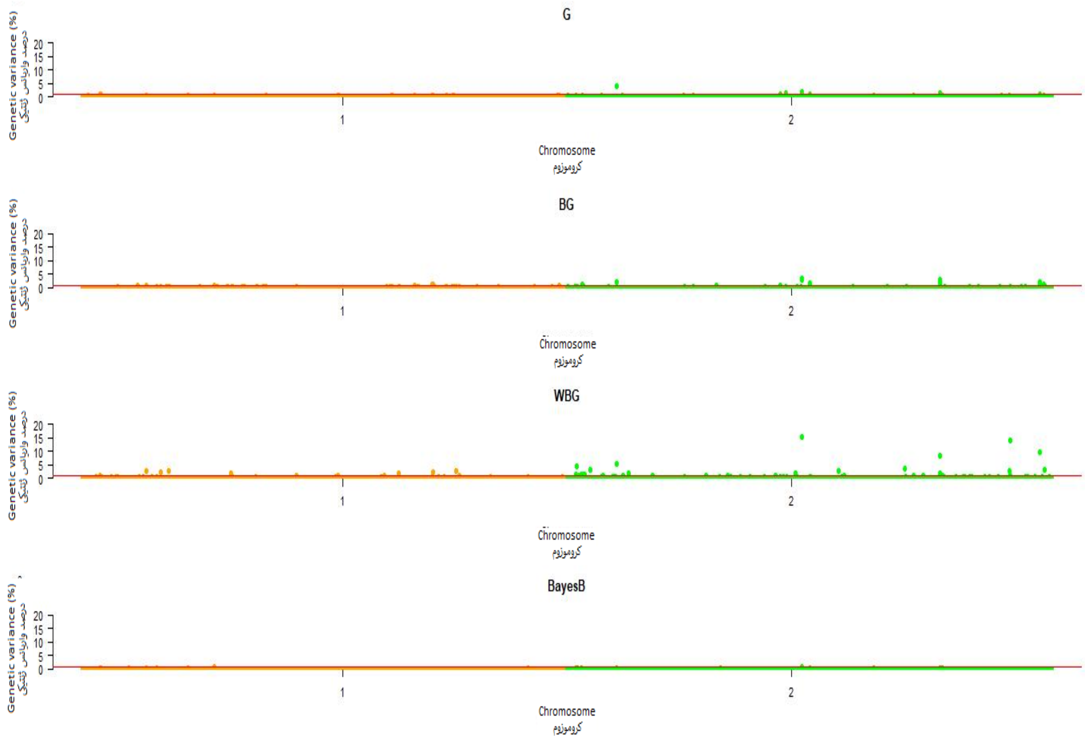
دارد و در GWAS چند نژادی QTL های با اثرات بزرگ که در این نواحی تحت انتخاب مشترک وجود دارند، شناسایی می‌شوند (van den Berg *et al*., 2016). نتایج پویش ژنومی نشان داد که استفاده از ماتریس WBG نسبت به سایر ماتریس‌های G و BG باعث شده است، که تعداد بیشتری نشانگر ژنومی مرتبط با QTLهای صفات T1 تا T4 شناسایی شوند (جدول ۳). حتی در مقایسه با بیز B نیز توان بالاتری برای پویش ژنومی چند جمعیتی داشت (شکل ۳) (Zhang et al., 2016). علاوه‌براین، نشانگرهایی وجود داشتند که بین ماتریس‌های BG و WBG مشترک شناسایی شدند، امّا در WBG از درصد واریانس توجیه شده بالاتری برخوردار بودند. به‌دلیل اینکه WBG در ابتدا نسبت به صفت اختصاصی می‌شود، می‌تواند با دقت بیشتری در وزن‌دهی‌های تکراری روش WssGWAS نشانگرهای اختصاصی یک صفت خاص با استفاده از اطلاعات چند نژاد را شناسایی نماید. همان‌طور که در شکل ۱ مشاهده شد، استفاده از ماتریس WBG باعث شد که پیش‌بینی‌های ارزش اصلاحی در اولین تکرار به همگرایی صحت در چرخه ۱۰ تکراری پویش ژنومی تک مرحله‌ای برسد. با استفاده از بیز B در صفات T1 تا T4 به‌ترتیب ۱، ۲، ۳ و ۵ نشانگر ژنومی با واریانس توجیه شده بالای یک درصد شناسایی شد، که این نشان می‌دهد روش بیز B به‌دلیل سختگیری‌ها و محدودیت‌هایی که در توزیع پیشین برآورد اثرات نشانگرهای ژنومی اعمال می‌کند، برای پویش ژنومی چند جمعیتی و پیش‌بینی ارزش اصلاحی ژنومی با جمعیت مرجع چند جمعیتی نسبت به روش‌های ssGBLUP و WssGBLUP مناسب نیست، امّا کمپر و همکاران (Kemper *et al*., 2015) نشان دادند که روش اماری بیز R نسبت به GBLUP برای نقشه‌یابی و پیش‌بینی ارزش اصلاحی به‌روش چند نژادی کارایی بهتری دارد. البته، مطالعه آن‌ها با استفاده از ماتریس G انجام شده است. در مطالعه اخیر نشان داده شد، ماتریس‌های BG و WBG نسبت به ماتریس G و بیز B توان بالاتری برای پیش‌بینی ارزش اصلاحی با WssGBLUP و پویش ژنومی چند نژادی دارند (شکل ۳). علاوه‌براین، مطالعات پویش ژنومی چند جمعیتی با روش آزمون رگرسیونی ساده نشان دادند که استفاده از اطلاعات فنوتیپی و ژنوتیپی چند نژاد گاو شیری باعث افزایش توان شناسایی نشانگرهای ژنومی مرتبط با QTLها می‌شود (Raven et al., 2014).

همچنین سانچز و همکاران (Sanchez et al., 2017) نشان دادند که استفاده از GWAS چند نژادی با استفاده از آزمون رگرسیونی ساده برای نقشه‌یابی QTLهای صفات مرتبط با تولید پروتئین شیر، توان روش آماری را برای شناسایی نشانگرهای ژنومی مرتبط با QTLها، به‌طور قابل ملاحظه‌ای بالا برده است. مطالعه دیگری توسط ون دنبرگ و همکاران (van den Berg et al., 2016) بر روی پنج نژاد گاو شیری نشان داد که GWAS چند نژادی نسبت به تک نژادی باعث بهبود دقت شناسایی نواحی QTLها می‌شود. مطالعات اخیر نشان دادند که استفاده از اطلاعات فنوتیپی افراد فاقد ژنوتیپ می‌تواند نقش مهمی در بهبود پویش ژنومی داشته باشد و علاوه‌بر آن، استفاده از ماتریس روابط خویشاوندی ژنومی با همبستگی بین جمعیتی بالاتر نسبت به سایر ماتریس‌های روابط خویشاوندی می‌تواند باعث بهبودی پویش ژنومی گردد. با توجه به مطالعات اخیر و نتایج این مطالعه، استفاده از جمعیت مرجع چند جمعیتی به‌دلیل افزایش مشاهدات فنوتیپی و اطلاعات ژنتیکی، باعث بهبود پویش ژنومی چند نژادی شده است. علاوه‌بر آن، استفاده از ماتریس‌های روابط خویشاوندی ژنومی بلوک‌بندی شده نظیر BG و WBG نسبت به ماتریس‌های روابط خویشاوندی ژنومی تک ساختاری نظیر G و GW باعث بهبود پویش ژنومی چندجمعیتی شده است.

جدول ۳- تعداد نشانگرهای ژنومی با بیش از یک درصد واریانس ژنتیکی افزایشی توجیه شده ماتریس‌های روابط خویشاوندی ژنومی

Table 3- Proportion of variance (%) explained of different genomic relationship matrices

صفات Traits	ماتریس روابط خویشاوندی ژنومی G		ماتریس روابط خویشاوندی ژنومی بلوک‌بندی شده Blok wise G		ماتریس روابط خویشاوندی ژنومی وزن‌دهی شده بلوک‌بندی Weighted Blok wise G	
	تعداد نشانگرهای ژنومی nSNPs	مجموع واریانس ژنومی توجیه شده (٪) ٪TV	تعداد نشانگرهای ژنومی nSNPs	مجموع واریانس ژنومی توجیه شده (٪) ٪TV	تعداد نشانگرهای ژنومی nSNPs	مجموع واریانس ژنومی توجیه شده (٪) ٪TV
T1	10±0.1	16.6±0.2	12±0.1	28.4±0.1	15±0.1	45.5±0.2
T2	17±0.2	33.8±0.1	17±0.1	36.3±0.1	24±0.1	68.1±0.1
T3	12±0.1	19.3±0.3	13±0.2	25.2±0.2	17±0.3	36.5±0.2
T4	18±0.3	14.4±0.1	20±0.2	31.5±0.1	27±0.1	59.3±0.1



شکل ۳- نمودارهای درصد واریانس‌های توجیه شده برای روش‌های مختلف

Figure 3- Manhattan plots of proportion of variance (%) explained of different methods

نتیجه‌گیری کلی

مطالعات اخیر و مطالعه حاضر نشان دادند که پویش ژنومی چند نژادی با روش آماری WssGWAS می‌تواند باعث بهبود پویش ژنومی چند نژادی برای شناسایی نواحی ژنومی مرتبط با صفات اقتصادی شود. علاوه‌بر آن، نتایج این مطالعه نشان داد که برای پویش ژنومی چند نژادی با روش آماری WssGWAS می‌توان به جای ماتریس روابط خویشاوندی G از ماتریس روابط خویشاوندی BG یا WBG استفاده نمود. ماتریس WBG به‌دلیل اینکه در ابتدای ورود چرخه تکراری وزن‌دهی روش WssGWAS توسط نتایج آنالیز روش آماری بیز B وزن‌دهی جداگانه شده است و هر نشانگر ژنومی در تکرار t=0 در ماتریس WBG واریانس مخصوص به خود را دارند با زمان کمتری به همگرایی صحت پیش‌بینی ارزش اصلاحی می رسند. بنابراین، پیشنهاد می‌گردد در مطالعات پویش ژنومی چند جمعیتی با استفاده از روش WssGWAS برای شناسایی نواحی ژنومی مرتبط با اثرات ژنتیکی افزایشی و غیر افزایشی صفات، از ماتریس‌های BG یا بهبود یافته آن WBG استفاده شود.

References

1. Aguilar , I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S., & Lawlor, T. J. (2010). Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science*, 93(2),743–752. https://doi.org/10.3168/jds.2009-2730
2. Alvarenga, A. B., Veroneze, R., Oliveira, H. R., Marques, D. B. D., Lopes, P. S., Silva, F. F., & Brito, L. F. (2020). Comparing alternative single-step GBLUP approaches and training population designs for genomic evaluation of crossbred animals. *Frontiers in Genetics*, 11(April),1–19. https://doi.org/10.3389/fgene.2020.00263
3. Campos, G. D. L., Vazquez, A. I., Fernando, R., Klimentidis, Y. C., Sorensen, D., de los Campos, G., Vazquez, A. I., Fernando, R., Klimentidis, Y. C., & Sorensen, D. (2013). Prediction of complex human traits using the genomic best linear unbiased predictor. *PLoS Genetics*, 9(7). https://doi.org/10.1371/journal.pgen.1003608
4. Cesarani, A., Lourenco, D., Tsuruta, S., Legarra, A., Nicolazzi, E. L., VanRaden, P. M., & Misztal, I. (2022). Multibreed genomic evaluation for production traits of dairy cattle in the United States using single-step genomic best linear unbiased predictor. *Journal of Dairy Science*, 105(6),5141–5152. https://doi.org/10.3168/jds.2021-21505
5. Cole, J. B., van Raden, P. M., O'Connell, J. R., van Tassell, C. P., Sonstegard, T. S., Schnabel, R. D., Taylor, J. F., & Wiggans, G. R. (2009). Distribution and location of genetic effects for dairy traits. *Journal of Dairy*

Science, 92(6),2931–2946. https://doi.org/10.3168/jds.2008-1762
6. Daetwyler, H. D., Kemper, K. E., van der Werf, J. H. J., & Hayes, B. J. (2012). Components of the accuracy of genomic prediction in a multi-breed sheep population. *Journal of Animal Science*, 90(10),3375–3384. https://doi.org/10.2527/jas.2011-4557
7. De Roos, A. P. W., Hayes, B. J., Spelman, R. J., & Goddard, M. E. (2008). Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. *Genetics*, 179(3),1503–1512. https://doi.org/10.1534/genetics.107.084301
8. Erbe, M., Hayes, B. J., Matukumalli, L. K., Goswami, S., Bowman, P. J., Reich, C. M., Mason, B. A., & Goddard, M. E. (2012). Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. *Journal of Dairy Science*, 95(7),4114–4129. https://doi.org/10.3168/jds.2011-5019
9. Gao, H., Christensen, O. F., Madsen, P., Nielsen, U. S., Zhang, Y., Lund, M. S., & Su, G. (2012). Comparison on genomic predictions using three GBLUP methods and two single-step blending methods in the Nordic Holstein population. *Genetics Selection Evolution*, 44(1),8. https://doi.org/10.1186/1297-9686-44-8
10. Goddard, M. E., & Hayes, B. J. (2009). Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nature Reviews Genetics*, 10(6),381–391. https://doi.org/10.1038/nrg2575
11. Gualdrón Duarte, J. L., Gori, A. S., Hubin, X., Lourenco, D., Charlier, C., Misztal, I., & Druet, T. (2020). Performances of Adaptive MultiBLUP, Bayesian regressions, and weighted-GBLUP approaches for genomic predictions in Belgian Blue beef cattle. *BMC Genomics*, 21(1). https://doi.org/10.1186/s12864-020-06921-3
12. Kemper, K. E., Reich, C. M., Bowman, P. J., Vander Jagt, C. J., Chamberlain, A. J., Mason, B. A., Hayes, B. J., & Goddard, M. E. (2015). Improved precision of QTL mapping using a nonlinear Bayesian method in a multi-breed population leads to greater accuracy of across-breed genomic predictions. *Genetics Selection Evolution*, 47(1),1–17. https://doi.org/10.1186/s12711-014-0074-4
13. Legarra, A., Christensen, O. F., Aguilar, I., & Misztal, I. (2014). Single Step, a general approach for genomic selection. *Livestock Science*, 166(1),54–65. https://doi.org/10.1016/j.livsci.2014.04.029
14. Legarra, A., Christensen, O. F., Vitezica, Z. G., Aguilar, I., & Misztal, I. (2015). Ancestral relationships using metafounders: Finite ancestral populations and across population relationships. *Genetics*, 200(2),455–468. https://doi.org/10.1534/genetics.115.177014
15. Liu, X., Tian, D., Li, C., Tang, B., Wang, Z., Zhang, R., Pan, Y., Wang, Y., Zou, D., Zhang, Z., & Song, S. (2023). GWAS Atlas: An updated knowledgebase integrating more curated associations in plants and animals. *Nucleic Acids Research*, 51(D1),D969–D976. https://doi.org/10.1093/nar/gkac924
16. Lourenco, D., Legarra, A., Tsuruta, S., Masuda, Y., Aguilar, I., & Misztal, I. (2020). Single-step genomic evaluations from theory to practice: using snp chips and sequence data in blupf90. *Genes*, 11(7),1–32. https://doi.org/10.3390/genes11070790
17. Lund, M. S., Sahana, G., de Koning, D. J., Su, G., & Carlborg, Ö. (2009). Comparison of analyses of the QTLMAS XII common dataset. I: Genomic selection. *BMC Proceedings*, 3(S1),1–8. https://doi.org/10.1186/1753-6561-3-s1-s1
18. Meuwissen, T. H., Solberg, T. R., Shepherd, R., & Woolliams, J. A. (2009). A fast algorithm for BayesB type of prediction of genome-wide estimates of genetic value. *Genetics Selection Evolution*, 41(2),1–10. https://doi.org/10.1186/1297-9686-41-2
19. Pedrosa, V. B., Boerman, J. P., Gloria, L. S., Chen, S.-Y., Montes, M. E., Doucette, J. S., & Brito, L. F. (2023). Genomic-based genetic parameters for milkability traits derived from automatic milking systems in North American Holstein cattle. *Journal of Dairy Science*, 106(4), 2613-2629. https://doi.org/10.3168/jds.2022-22515
20. Pérez, P., & De Los Campos, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics*, 198(2),483–495. https://doi.org/10.1534/genetics.114.164442
21. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., De Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81(3),559–575. https://doi.org/10.1086/519795
22. Raven, L. A., Cocks, B. G., & Hayes, B. J. (2014). Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. *BMC Genomics*, 15(1). https://doi.org/10.1186/1471-2164-15-62
23. Reales, G., & Wallace, C. (2023). Sharing GWAS summary statistics results in more citations. *Communications Biology*, 6(1),6–11. https://doi.org/10.1038/s42003-023-04497-8
24. Sanchez, M. P., Govignon-Gion, A., Croiseau, P., Fritz, S., Hozé, C., Miranda, G., Martin, P., Barbat-Leterrier, A., Letaïef, R., Rocha, D., Brochard, M., Boussaha, M., & Boichard, D. (2017). Within-breed and multi-breed GWAS on imputed whole-genome sequence variants reveal candidate mutations affecting milk protein composition in dairy cattle. *Genetics Selection Evolution*, 49(1),68. https://doi.org/10.1186/s12711-017-0344-z
25. Su, G., Christensen, O. F., Janss, L., & Lund, M. S. (2014). Comparison of genomic predictions using genomic

relationship matrices built with different weighting factors to account for locus-specific variances. *Journal of Dairy Science*, 97,6547–6559. https://doi.org/10.3168/jds.2014-8210
26. Su, G., Guldbrandtsen, B., Gregersen, V. R., & Lund, M. S. (2010). Preliminary investigation on reliability of genomic estimated breeding values in the Danish Holstein population. *Journal of Dairy Science*, 93(3),1175–1183. https://doi.org/10.3168/jds.2009-2192
27. Tiezzi, F., & Maltecca, C. (2015). Accounting for trait architecture in genomic predictions of US Holstein cattle using a weighted realized relationship matrix. *Genetics Selection Evolution*, 47(1),1–13. https://doi.org/10.1186/s12711-015-0100-1
28. van den Berg, I., Boichard, D., & Lund, M. S. (2016). Comparing power and precision of within-breed and multibreed genome-wide association studies of production traits using whole-genome sequence data for 5 French and Danish dairy cattle breeds. *Journal of Dairy Science*, 99(11),8932–8945. https://doi.org/10.3168/jds.2016-11073
29. Vanraden, P. M. (2008). Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science*, 91(11),4414–4423. https://doi.org/10.3168/jds.2007-0980
30. Veroneze, R., Lopes, P. S., Lopes, M. S., Hidalgo, A. M., Guimarães, S. E. F., Harlizius, B., Knol, E. F., van Arendonk, J. A. M., Silva, F. F., & Bastiaansen, J. W. M. (2016). Accounting for genetic architecture in single- and multipopulation genomic prediction using weights from genomewide association studies in pigs. *Journal of Animal Breeding and Genetics*, 133(3),187–196. https://doi.org/10.1111/jbg.12202
31. Wang, H., Misztal, I., Aguilar, I., Legarra, A., Fernando, R. L., Vitezica, Z., Okimoto, R., Wing, T., Hawken, R., & Muir, W. M. (2014). Genome-wide association mapping including phenotypes from relatives without genotypes in a single-step (ssGWAS) for 6-week body weight in broiler chickens. *Frontiers in Genetics*, 5(MAY),134. https://doi.org/10.3389/fgene.2014.00134
32. Wang, H., Misztal, I., Aguilar, I., Legarra, A., & Muir, W. M. (2012). Genome-wide association mapping including phenotypes from relatives without genotypes. *Genetics Research*, 94(2),73–83. https://doi.org/10.1017/S0016672312000274
33. Wientjes, Y. C. J., Bijma, P., Vandenplas, J., & Calus, M. P. L. (2017). Multi-population genomic relationships for estimating current genetic variances within and genetic correlations between populations. *Genetics*, 207(2),503–515. https://doi.org/10.1534/genetics.117.300152
34. Wientjes, Y. C. J., Calus, M. P. L., Goddard, M. E., & Hayes, B. J. (2015). Impact of QTL properties on the accuracy of multi-breed genomic prediction. *Genetics Selection Evolution*, 47(1),42. https://doi.org/10.1186/s12711-015-0124-6
35. Wijesena, H. R., Nonneman, D. J., Snelling, W. M., Rohrer, G. A., Keel, B. N., & Lents, C. A. (2023). gBLUP-GWAS identifies candidate genes, signaling pathways, and putative functional polymorphisms for age at puberty in gilts. *Journal of Animal Science*. https://doi.org/10.1093/jas/skad063
36. Wu, X., Lund, M. S., Sun, D., Zhang, Q., & Su, G. (2015). Impact of relationships between test and training animals and among training animals on reliability of genomic prediction. *Journal of Animal Breeding and Genetics*, 132(5),366–375. https://doi.org/10.1111/jbg.12165
37. Zhang, X., Lourenco, D., Aguilar, I., Legarra, A., & Misztal, I. (2016). Weighting strategies for single-step genomic BLUP: An iterative approach for accurate calculation of GEBV and GWAS. *Frontiers in Genetics*, 7(AUG),151. https://doi.org/10.3389/fgene.2016.00151
38. Zhang, Z., Liu, J., Ding, X., Bijma, P., de Koning, D. J., & Zhang, Q. (2010). Best linear unbiased prediction of genomic breeding values using a trait-specific marker-derived relationship matrix. *PLoS ONE*, 5(9),1–8. https://doi.org/10.1371/journal.pone.0012648
39. Zhang, Z., Ober, U., Erbe, M., Zhang, H., Gao, N., He, J., Li, J., & Simianer, H. (2014). Improving the Accuracy of Whole Genome Prediction for Complex Traits Using the Results of Genome Wide Association Studies. *PLoS ONE*, 9(3),1–12. https://doi.org/10.1371/journal.pone.0093017
40. Zhou, L., Lund, M. S., Wang, Y., & Su, G. (2014). Genomic predictions across Nordic Holstein and Nordic Red using the genomic best linear unbiased prediction model with different genomic relationship matrices. *Journal of Animal Breeding and Genetics*, 131(4). https://doi.org/10.1111/jbg.12089